

E

V 8078

The Display of Three-Dimensional Video Images

A. R. L. TRAVIS

XP-000755850

p.

1817-1832 =

(16)

Invited Paper

p.d. 11-1997

Three-dimensional images can be pixellated in three distinct ways: volumetric, holographic, and autostereoscopic. The latter excels if images of opaque objects are to be displayed with wide fields of view, and the quality of view-sequential displays with 1° per view now appears adequate for general application. Although in principle autostereoscopic pixellation gives a true three-dimensional image, $1/10^\circ$ per view is needed to avoid flaws in a typical display. This approximately equals the diffraction limit, and the information content is no less than that of a hologram. A hybrid of holographic views and view-sequential multiplexing promises images with the field of view of autostereoscopic images but the significantly greater resolution and depth of holograms. Light valves and high-frame-rate arrays already have the space-bandwidth product needed to display such images, and further advances in photonic switches and gigahertz telecommunications look set to promote the display of such high-quality three-dimensional video images.

Keywords—Displays, holographic, television, 3-D, three-dimensional, video.

I. INTRODUCTION

Conventionally televised images are two dimensional (2-D) yet enable sufficient depth perception that surgeons, for example, are able to operate by them. Nevertheless, when depth perception is critical, as it is in manipulative activities like surgery, depth perception is quicker and more reliable if the images have a three-dimensional (3-D) content [1]. Television and video games are likely to be more realistic with three-dimensional images, and the human interpretation of complicated visual data more simple. There has therefore been a renewed interest in three-dimensional television, and while there has been detailed work [2] on the systems needed for this, a concentrated analysis is desirable of the component on whose radical evolution the rest of the system will depend: the display.

It can come as a surprise to learn how little is needed to make a display for crude three-dimensional images. For example, one need merely take the liquid crystal display from a typical laptop computer, swap the back illuminator for a lens, and place a spot source of light some distance

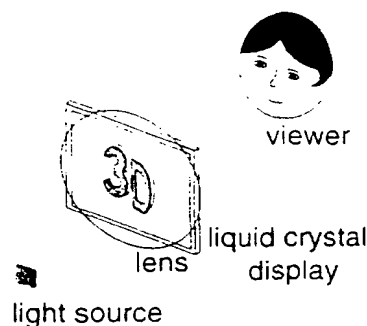


Fig. 1. One can display a three-dimensional image by showing views of the object on a liquid crystal display and illuminating each to an appropriate direction.

behind the lens, as shown in Fig. 1. The spot source might comprise a laser beam incident on a translucent screen, which under the action of the lens will illuminate the display with rays that converge to form an image of the source. Since the picture on the display will be visible only if observed from within the confines of the image of the source, the picture will have a restricted field of view. It is set up to be one view of a three-dimensional object. Other views of the three-dimensional object can be made visible to other areas by deflecting the laser beam to a different position for each view. If this operation is repeated at a rate sufficient to avoid flicker, and if the whole of the plane of convergence is illuminated, then the result will be a steady three-dimensional image even if, as will be shown later, the display is viewed away from the plane of convergence.

The three-dimensional images formed by such a display will be crude because both the amorphous silicon transistors and the nematic liquid crystal typically found in a liquid crystal display switch too slowly to form more than one clear view. Furthermore, the size of the display will be limited by the complexity of liquid crystal displays, with the manufacture of large devices at present expensive.

This example conveniently illustrates some features of creating three-dimensional displays. Because of the extra dimension, high-quality three-dimensional images require data for one to three orders of magnitude more pixels than two-dimensional images, dependent on resolution and the

Manuscript received June 18, 1997; revised August 18, 1997.

The author is with the Department of Engineering, Cambridge University, Cambridge CB2 1PZ U.K.

Publisher Item Identifier S 0018-9219(97)08216-9.

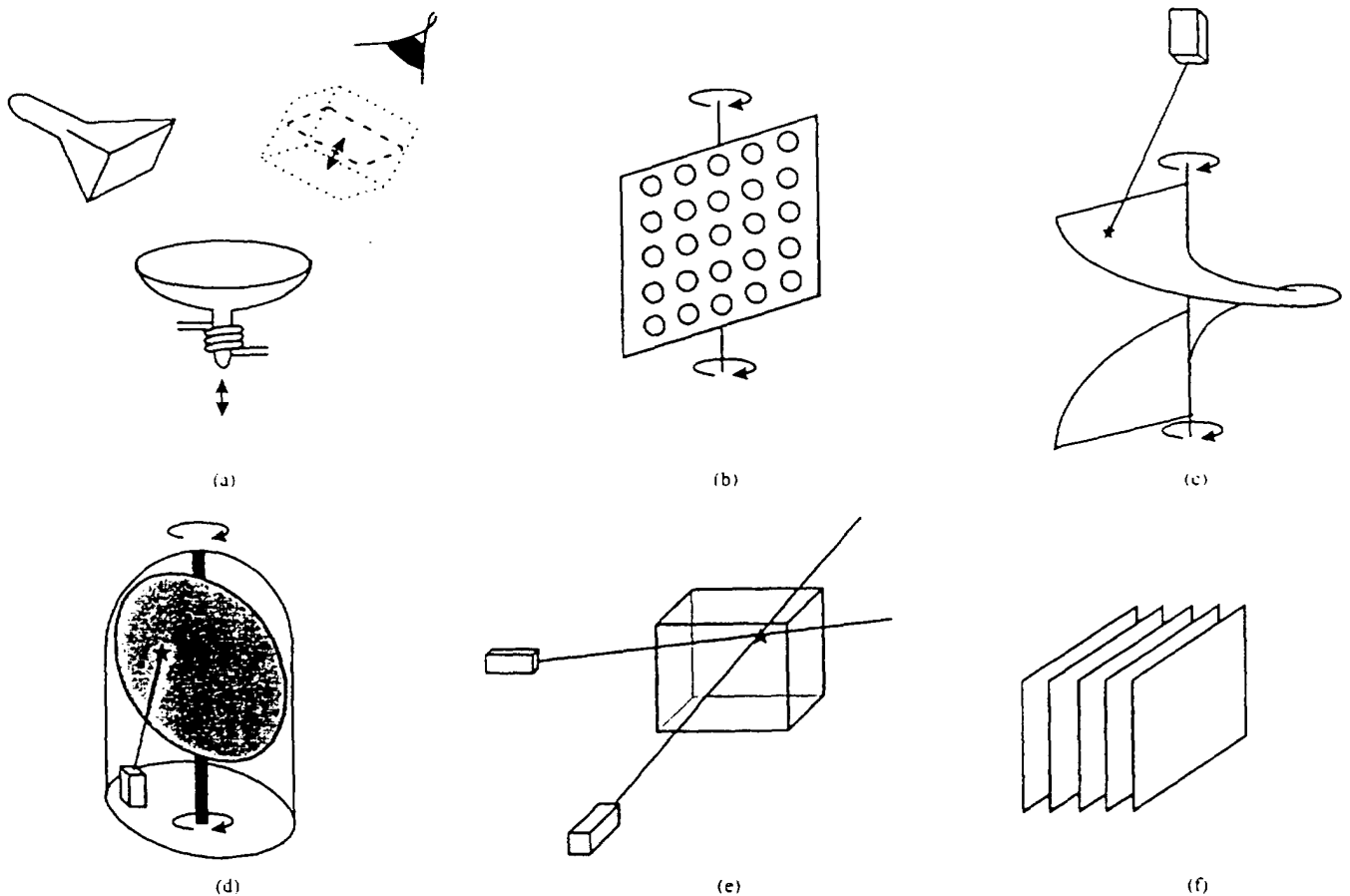


Fig. 2. Volumetric 3-D displays. (a) Vibrating mirror. (b) Spinning LED's. (c) Spinning translucent screen. (d) Spinning phosphor disc. (e) Two-photon absorption. (f) Stacked LCD's.

field of view. The first challenge faced by the designer is physically to distribute these data across the display's screen at a sufficient rate. The second challenge is that of providing the screen itself with a sufficient space-bandwidth product, i.e., enough pixels each switching sufficiently quickly to transfer the data into modulated light. Last is the challenge of enabling the manufacture of a display with these properties without requiring prohibitive precision or cost.

Each of these challenges is familiar to the designer of displays for high-definition two-dimensional images, and it is arguable that, pixellation and optics aside, three-dimensional video images are merely a technological extension of their two-dimensional predecessors. Nevertheless, the variety of schemes recently put forward is bewildering, so this paper will proceed to review some of the more successful technologies and show that they comprise three distinct schemes: volumetric, holographic, and autostereoscopic. One of these, autostereoscopic, shows promise but contains unwanted lines (flaws) at low resolution, so this paper goes on to quantify this. Section IV considers what resolution is needed for an autostereoscopic three-dimensional image to be free of such flaws and shows that for a typical size of display, one can do as well if not better with a hologram. Section V proposes a hy-

brid of autostereoscopic and holographic pixellation, which gives the advantages of both. The sixth section shows how photonic devices make the display of such images possible. This paper concludes by evaluating the bandwidth of the latest photonic devices, noting the trend in three-dimensional display toward the integration of the display with the computer and the future dependence of both on advances in optical switches operating at gigahertz rates.

II. A REVIEW OF THREE-DIMENSIONAL VIDEO DISPLAYS

A. Volumetric Displays

One way of screening a three-dimensional image is to extend the principle of conventional television to the third dimension by making a device capable of emitting light at any point in a volume (Fig. 2). Perhaps the earliest way of doing this was to reflect light from a cathode ray tube off a circular mirror that vibrated like a loudspeaker [3], [4]. An image of the cathode ray tube formed at varying distances from the mirror, thereby sweeping out a three-dimensional volume, but the supporting structure was heavy and the field of view limited. Light emitting diode screens [5], [6] or laser-scanned displays [7], [8] have been used instead of

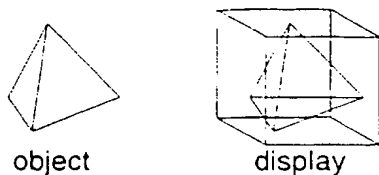


Fig. 3. A three-dimensional array of light emitters cannot display opaque images.

a cathode ray tube, but the mechanism for scanning depth remains cumbersome.

An almost unrestricted field of view can be provided by spinning a two-dimensional array of light emitters through a three-dimensional volume. Among other ways, this has been done with an array of light emitting diodes [9], a translucent screen that is scanned by lasers [10]–[12], and a phosphor screen that is scanned (inside a vacuum) by electron beams [13]. The last way has the advantage of a cheap screen and scanning mechanism, but any rotating screen system has a singularity at the axis of rotation.

An unrestricted field of view without an axis of singularity can be provided by scanning a pair of laser beams across a transparent material, which emits light isotropically where the laser beams intersect [14], [15]. An image of approximately one cubic centimeter has been demonstrated, but even a larger image would, like both vibrating-mirror and spinning-disc displays, provide only for the emission of light and not for its absorption. Each of these displays is therefore not able to provide opacity, so while the displayed images are three dimensional, they are necessarily translucent (Fig. 3). Schemes for the display of opaque images have been proposed—for example, stacking liquid crystal displays into a volume [16]—but even if these were interleaved with light emitters, the result would still be incapable of displaying reflections or specularity.

The advantage of volumetric displays is that they can provide an unrestricted field of view without excessive data rates. This means that volumetric displays are not without potential application, for example, in air-traffic control or battle management. But the ideal is a device free of all optical restrictions. One device that can display any three-dimensional image with certainty is the hologram.

B. Holographic Displays

A hologram effectively freezes the optical wave fronts scattered off a three-dimensional object by recording their complex amplitude. Dynamic holograms are often proposed as a way of displaying a three-dimensional image [17] (Fig. 4). A gray-scale hologram is merely a high-resolution two-dimensional image, and conventional liquid crystal displays can be used to display such a hologram, albeit with a narrow (4° at present) field of view [18], [19]. Wider fields of view require pixellation too fine for active matrix displays, but ferroelectric liquid crystal displays can be made with several thousand columns at realistic yields [20]. Nevertheless, this leaves the problem of bonding to several thousand connectors, which one scheme avoids by scanning the back of an optically addressed liquid crystal

display with a cathode ray tube [21], [22]. Even with this improvement, the resolution of any liquid crystal display cannot be less than two or three times the cell gap, the result of which is to restrict the field of view of the hologram to a few degrees.

Acoustooptic modulators provide phase modulation and have been used to display color dynamic holograms [23], [24]. The difficulties with scanning mirrors and bulk optics can to some extent be avoided [25], but once again, there is also the difficulty of modulating light at a resolution sufficient to get a wide field of view. For a 20° field of view, a resolution of approximately $2.5 \mu\text{m}$ is needed, which, with the speed of sound in a typical acoustooptic crystal being $5 \text{ km}\cdot\text{s}^{-1}$, requires acoustic modulation at 2 GHz. Because of the attenuation at such frequencies, the crystal must be kept smaller than the average display. Even if attenuation could be avoided, the data rates are too high for ease of operation. Operation can be eased by synthesizing the hologram within the crystal from a number of independently modulated frequencies, but unless the phase of these frequency constituents is actively controlled, the result is not a holographic reproduction of the original three-dimensional image.

It is because of the need to reproduce optical phase that the data rates of a true holographic display are extreme. But the human eye is no more sensitive to the phase of a three-dimensional image than it is to the complete optical spectrum of a color image. Just as color images need comprise only red, green, and blue primaries, a three-dimensional image need comprise only the correct distribution of ray intensity versus position and direction that is specified by autostereoscopic pixellation.

C. Autostereoscopic Displays

Autostereoscopic displays are named to distinguish them from their stereoscopic predecessors, which require the user to wear spectacles. Stereoscopic displays require two separate views to be generated and then presented, one to each eye. Among the latest stereoscopic displays, one has spectacles comprising a pair of liquid crystal shutters that are synchronized to a screen that displays alternate left- and right-eye pictures; with a sufficiently high frame rate, the viewer sees a flicker-free image. The image provides stereopsis, i.e., the binocular perception of depth, but not kineopsis, which is the monocular perception of depth we accumulate by subconsciously moving our heads around a scene. Of the two, stereopsis is confined mainly to animals such as predators and primates who need to make instant estimates of depth, and it is arguable that even in these species, kineopsis is a more relied-upon determinant of depth in static situations. Viewers can experience nausea after prolonged viewing of stereoscopic displays [26], which may be due to subconscious awareness of the lack of kineopsis, but the real problem with stereoscopic displays is that spectacles get lost.

Spectacles become unnecessary if each view is projected into one eye, which can be done using the display described in Section I. Such displays are called autostereoscopic

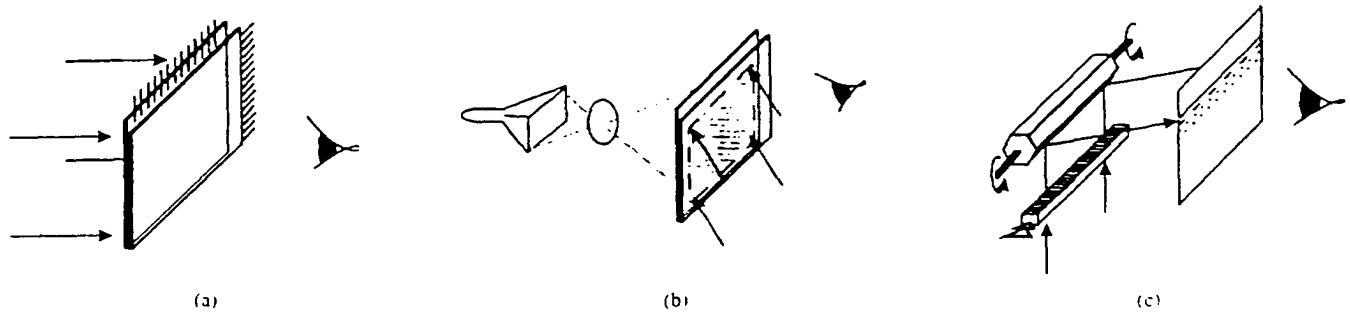


Fig. 4. Holographic 3-D displays. (a) Hologram on an LCD. (b) Hologram on an OASLM. (c) Acoustooptic hologram.

(Fig. 5), a word that, like television, is an unhappy mix of Greek and Latin that seems peculiar to the display industry. One cannot expect a viewer to keep his head fixed merely for the convenience of the display, so one approach that is attracting great interest is to continually monitor the position of the viewers' heads and adjust the projection optics and visual content accordingly.

More than one pair of eyes can in principle be tracked, and if the content of each view is matched to eye position, then the display can provide for both stereopsis and kineopsis. Furthermore, it might be possible to display views with approximately the right perspective by guessing the distance of each viewer from the screen through measurement of the distance between their eyes, so that almost the only missing depth cue would be accommodation (the ability to focus on off-screen pixels). Although the resulting image would therefore be something short of truly three dimensional, it is unlikely that viewers would notice.

The design of the display for such systems is relatively straightforward because the data rate of conventional video needs to be increased only by a factor of two (or four for two viewers, etc.) [27], [28], so that the major challenge becomes that of identifying and tracking the viewers. Demonstrators have been built that require the viewer to wear an infrared reflecting spot [29] or a magnetic sensor [30]–[32], but many authors are coy on their plans for tracking bare heads [33].

An elegant approach is to side-illuminate the head with infrared light so that one eye is illuminated and the other in shadow [34], [35], but the shadows of more than one viewer can fall on each other. Another impressive approach is to track the hair/face boundary of viewers [36], while a system that tracks the eye, nose, and lips of a face has achieved 80% reliability with the face of the designer [37]. But the latter is slow, tracks only one face, and is less effective with a variety of faces. Advances in technologies like speech, handwriting, and object recognition mean that the day must surely come when systems will be aware of their surroundings, but the development of such machine intelligence will herald a new generation of computing, and progress in these areas so far has been slow. Meanwhile, the possibility of irritating glitches due to intermittently unfamiliar situations is never quite excluded, and users are notoriously intolerant of such weaknesses.

Multiple-view autostereoscopy makes the position of the viewers' heads irrelevant because the display projects views to every position where a viewer might be. It will be left until the next section to convince the skeptical that such an image can be truly three dimensional, but with the need for a many-fold increase in bandwidth, the design of the display now becomes daunting.

The lenslet array is perhaps the longest established such autostereoscopic technology [38]–[41], first developed to give three-dimensional photography and now being applied to displays. Each lenslet occupies the area that would be taken up by a single pixel if the display were configured for two-dimensional images, and underneath the lenslet is a series of subpixels (one for each view) whose emissions are collimated by the lenslet to the appropriate direction. Although lenslets magnify the dead zone between adjacent subpixels, this can be smoothed out [42], but the numerical aperture of simple lenslets restricts the field of view of lenslet displays to a total angle of approximately 15° . Outside this angle, the three-dimensional image repeats itself, which can be irritating.

If an array of diffraction gratings is used instead of an array of lenslets, it is possible to get wider fields of view without dead zones or repeating views [43]–[45], but both grating and lenslet array displays require an underlying display whose resolution is the product of the resolution of each view and the number of views: a substantial manufacturing challenge. Nevertheless, high-resolution displays are in prospect, and the latest lenslet array displays assembled in laboratories have eight views at color video graphics adaptor (VGA) resolution.

High manufacturing yields are unnecessary if one makes a display by lining up several video projectors behind a lens [46], [47]. In this system, the projectors image one view each onto the lens, and the lens makes each view visible to a different direction. The projectors must be precisely aligned and have uniform brightness, and the projection lenses must be carefully designed to adjoin one another without perceptible gaps.

Both lenslet arrays and multiprojector systems multiplex the views of a three-dimensional image from spatially distinct subpixels, but one can also use the persistence of human vision to multiplex video images over time. It is possible to take what amounts to a single lenslet with

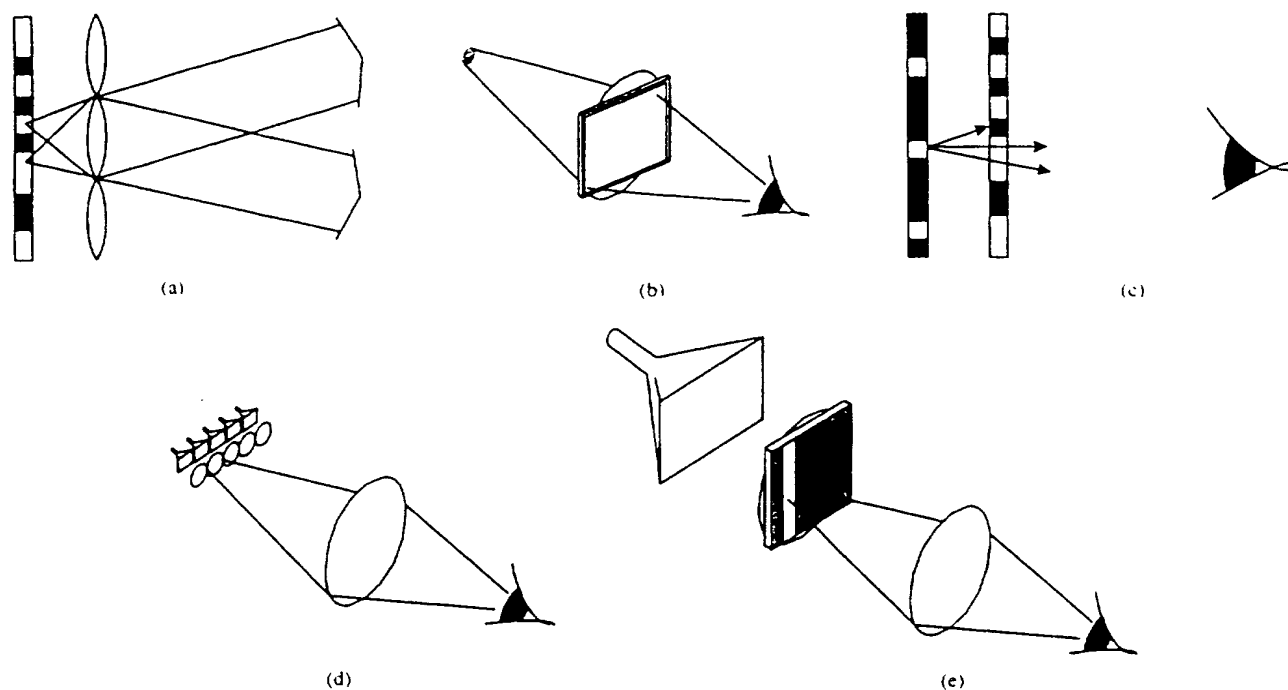


Fig. 5. Autostereoscopic 3-D displays. (a) The lenticular array. (b) Parallel-illuminated LCD. (c) Line-illuminated LCD. (d) Multiple projectors. (e) Shuttered projector.

subpixels from a lenslet array display and raster scan it across a screen with spinning mirrors [48], but it is difficult to see how to multiplex across the whole screen in this way without moving parts.

The alternative is to multiplex the views over time, and with the lenslet array display, this can be done by replacing the lenslet array with a low-resolution array of slits [49], [50]. Due to pin-hole optics, the slits act at any instant like lenslets, and with a low-resolution display underneath produce a low-resolution three-dimensional image. By scanning the slits over the underlying display, it is possible to time multiplex the equivalent of a full-resolution lenslet array but with no lens aberration and no need for high-resolution subpixels. Slits, however, waste light; a less wasteful method of getting the same optical effect is to exchange the slit for line illumination [51], [52]. Similar but perhaps less complex is the time-multiplexed concept described in Section I [53], [54]. Both latter approaches have the great virtue of wasting no more light than a conventional liquid crystal display, but both require a transmissive spatial light modulator with a high frame rate.

Polycrystalline silicon transistors and ferroelectric liquid crystals each switch an order of magnitude faster than their amorphous silicon and nematic predecessors. Using these, a small liquid crystal display with a frame rate of 1 kHz has been demonstrated [55]. Cadmium selenide and amorphous diamond transistors also switch quickly, and fast-switching gray-scale modulation is made possible by the distorted helix and electroclinic effects, by monostable or domain-switching ferroelectric liquid crystals, and by antiferroelectric liquid crystals. Great resources were needed to develop even the existing liquid crystal displays,

however, and greater confidence in the desirability of video three-dimensional images will be needed before advanced liquid crystal displays are developed.

A time-multiplexed cousin of the multiprojector system can be constructed by replacing the several projectors with a single large projector, whose projection lens covers the whole area filled by the multiple projectors [56], and placing over the lens a mechanical [57]–[59] or liquid crystal [60] shutter that blocks light from all but one area. At any instant, the projector does the same as one of the projectors in the spatially multiplexed system, but at successive instants, different areas of the shutter are made transparent so that each view of the three-dimensional image can be projected in turn. Careful alignment is unnecessary, so a cathode ray tube can be used without the expense of beam indexing. Indeed, the concept is so fault tolerant that the author was able to assemble a crude system from a cheap video display unit and a couple of fresnel lenses.

In the contest between spatial and time multiplexing, it is the spatially multiplexed lenslet array that seems to be receiving the most attention from manufacturers, perhaps because the high-resolution yields that are required present a manufacturing challenge of a kind that manufacturers have faced so successfully in the past. Certainly, the history of the semiconductor industry has seen inexorable increases in resolution, but there also have been increases in switching speed of a similar magnitude, to those that will be needed for time-multiplexed three-dimensional video.

It is arguably time multiplexing that has allowed the cathode ray tube to dominate the display of two-dimensional video, and a time-multiplexed 3-D projection system using cathode ray tubes produced an image comprising eight

monochrome VGA views several years ago. Despite being bulky and optically inefficient, this system is robust and flexible and continues to use the high data rate of cathode ray tubes to produce image qualities in advance of lenslet arrays. It is perhaps all the more remarkable that a crude concept with many similarities was built more than fifty years ago by Baird [61], [62].

The latest autostereoscopic displays produce images in which each view is visible across an arc of 1° , and there is a consensus among those who have seen such images that for the first generation of this technology, 1° per view will suffice. Experience with two-dimensional video has shown that expectations of resolution invariably increase, but Section IV will show that for a VGA picture, there is no point having an angle per view finer than 0.1° . It follows that for subsequent generations of 3-D display, the angle per view for VGA resolution pictures will be somewhere between 0.1° and 1° , depending on perception and cost.

If a display is to produce true three-dimensional images, then it should be able to project the image of pixels at various depths, and the viewer should see perspective that changes with their distance from the image. While it is clear that volumetric and holographic displays can do this, the description of autostereoscopic pixellation so far provided makes it less apparent that autostereoscopic displays can also project true three-dimensional images. The next section aims to remedy this.

III. COARSE AUTOSTEREOSCOPIC PIXELLATION

With the first concepts for television it was proposed to use systems of spinning slits, and it is instructive to consider what happens if a spinning slit is placed in front of the hologram of a three-dimensional image. It is a matter for simple experiment to look at a three-dimensional object through a spinning slit, and it is observed that the scene is unchanged except for being dimmer and perhaps slightly blurred. A hologram should reproduce the wave fronts of a monochromatic three-dimensional image exactly, so a hologram seen through a spinning slit should also appear unchanged. What makes such an experiment significant is that the slit prevents superposition between light from areas of the hologram alternately exposed by the slit. So we can consider the hologram as an assemblage of independent slit-sized subholograms. The results of this experiment would be no different if a raster scanning hole were used instead of a spinning slit, so a hologram can be further considered as a two-dimensional array of hole-sized subholograms.

The subholograms are different from the pixels of a two-dimensional image in that the intensity of light is a function of direction from which the subhologram is observed, as well as a function of the subhologram's position. Since there are two coordinates of direction (azimuth and elevation) as well as two coordinates of position, a system of four real coordinates is required for the true reproduction of a three-dimensional image.

Now imagine that a second spinning slit is placed some distance away from the first, as shown in Fig. 6, and that it

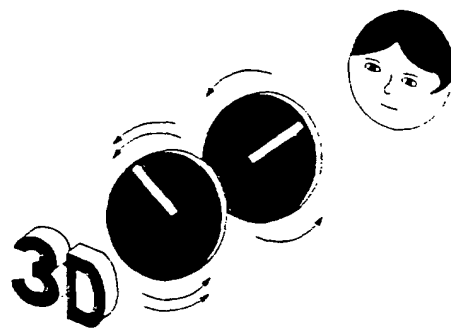


Fig. 6. A three-dimensional object looks the same when seen through a pair of slits, one spinning rapidly and the other slowly.

spins sufficiently quickly that there is no Moiré fringing observed between the two. We would expect the three-dimensional scene to remain unchanged except for being yet dimmer and perhaps more blurred. Only light traveling from the first slit through the second will be exposed at a single instant, and if both slits are replaced by raster scanning holes of sufficiently small diameter, then the light passing through both will necessarily approximate to a single Gaussian ray. Because the second hole exposes rays traveling to different directions alternately, it removes superposition between the rays. It follows that even if it uses entirely incoherent light, a system that modulates rays as a function of both position and direction will suffice to display a true three-dimensional image.

This thought experiment demonstrates that autostereoscopic displays have the potential to produce true three-dimensional images, but the images will only be genuinely three dimensional if they comprise enough views, and the eight or so views available from existing autostereoscopic displays are too few. If the image is not genuinely three dimensional, how different does it look?

Taking the display described in Section I as our model of explanation, imagine as before that the observer looks at a liquid crystal display in front of a lens but that the spot source of light behind the lens is created by illuminating one of an array of abutting light sources in the lens' focal plane. The image on a liquid crystal display is only entirely visible to an eye if at all points, the display is illuminated by rays of light that travel toward the eye. So the eye will only see the whole of one view if the eye is far from the display such as to subtend approximately the same angle to all points on the display.

If, however, the eye moves closer to the display, it will subtend an increasingly different angle to one side of the display than the other. In the first instance, the screen will divide into two zones, one illuminated by one element of the array and one by that next to it. Since a different view appears with each element of the array, the image accumulated over time by the eye will be the equivalent of cutting the left half off one view and the right half off the other and sticking the two together. If the angle between views is fine, then the two views are similar enough that the edge between the two halves is unnoticeable. But if the

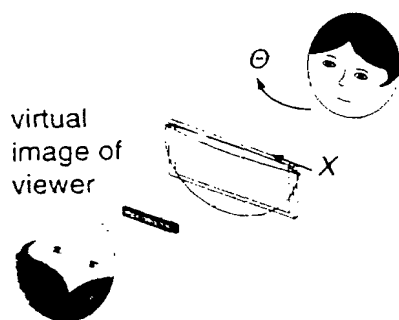


Fig. 7. Close to the screen, a perspective image is seen whose composition can be determined by ray tracing.

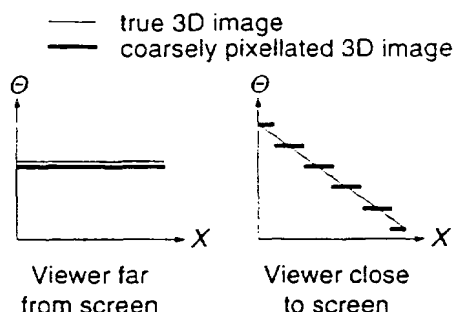


Fig. 8. A plot of pixel direction θ versus pixel position X can be used to identify what the viewer sees on the display, no matter how distant it is.

spacing between views is coarse, then the content of each view differs markedly from that next to it, so that at the boundaries between halves, there is an image discontinuity that looks like a flaw line.

Moving the eye closer still, the precise composition of the image seen can be determined by using the optical trick of tracing rays backward, as shown in Fig. 7. Starting with all those rays that reach the eye's pupil, one can trace their paths backward through the liquid crystal display and lens to the illuminators and, ignoring these, to a point where the rays all converge. All rays reaching the eye can be imagined to originate from this point so that it constitutes a virtual image of the eye's pupil. Lines drawn from this point through the edges of each element of the array intersect with the liquid crystal display to delineate the area of the liquid crystal display that is made visible by that element.

It is the coarseness of view spacing that causes flaw lines rather than autostereoscopy itself because the process described above is exactly how one gets perspective with a real object. Far from the object, an eye will see a view comprising a parallel projection of the object in that direction, but close up, the eye will subtend a different angle to one side of the object than to the other. Therefore, the eye will see rays from one side of the object that are part of a different parallel projection from rays from the other. Fig. 8 plots typical results showing schematically how the coarsely pixellated composition seen by the eye compares with the true 3-D image.

The effects of Fig. 7 can be demonstrated on an autostereoscopic display by configuring each view as a hor-

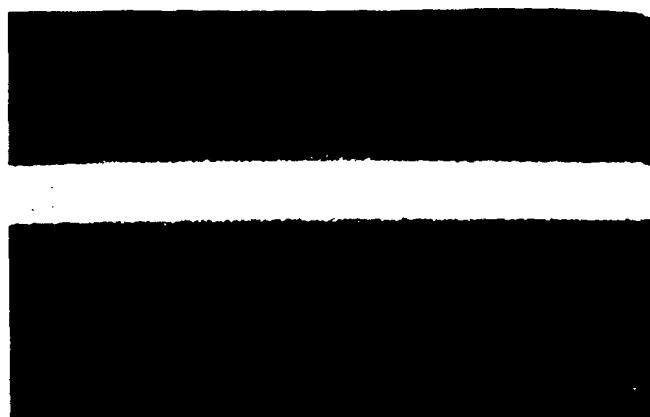


Fig. 9. Distant photograph of an autostereoscopic display on which each view comprises a horizontal bar. (Permission for reprint, courtesy Society for Information Display.)

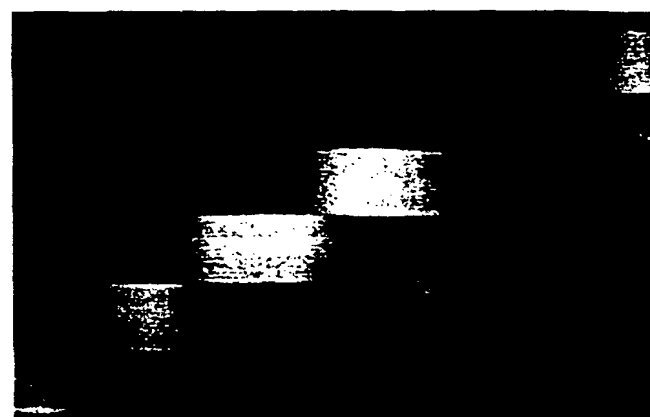


Fig. 10. Close-up photograph of an autostereoscopic display on which each view comprises a horizontal bar. (Permission for reprint, courtesy Society for Information Display.)

izontal bar: the left-most view with the bar at the top of the screen, the right-most with the bar at the bottom, and the remainder spaced evenly between. Rather conveniently, the result is the synthesis of the diagram of Fig. 8: far from the display, a single view—i.e., a single horizontal bar—is visible, whereas close to the display, visible parts of different views comprise a staircase of bars, as shown in Figs. 9 and 10.

It is tempting to suppose that the 3-D image generated by a shuttered cathode ray tube would be smoother if the scanning shutter was scanned continuously as each view was written on the cathode ray tube instead of being moved by a single shutter width between the display of each view [63]. The idea is that this continuous movement of the shutter might smooth discontinuities between adjacent views, which arise when the angles between them are too coarse. Considering only the horizontal dimension, assume that as the cathode ray tube traces out the X coordinate, the shutter gradually moves by one shutter width. This gives a gradual change of θ with X , so the pixellation in the X/θ diagram is slanted. There will be a distance from the display where someone looking at it will see a picture that can be represented on the diagram by a diagonal that is

parallel to the pixellation lines, so that at this distance, the viewer will see a single view. This is exactly the result we would get if we put a weak lens in front of the screen of a conventional autostereoscopic display, so the 3-D image is not smoothed but merely distorted.

Tolerant as the eye is of flaw lines between views, they nevertheless remain apparent. The claim that an autostereoscopic display produces a true three-dimensional image can only be valid if the spacing between views is sufficiently fine; but just how fine is sufficient?

IV. 3-D PIXELLATION

The spacing of 1° per view that was reported in Section II to be satisfactory for the present generation of displays requires 60 views for a typical field of view of 60° . It is tempting to state that flaws will only be eliminated on an autostereoscopic display if views are as finely separated as the human eye can resolve [64], but it was one of the breakthroughs in the development of two-dimensional video to realize that such detail is unnecessary. This section assumes that a three-dimensional image will be acceptable if with the same pixel dimensions as the equivalent two-dimensional image, it can be displayed without flaws.

The volumetric array is the format in which computer-aided design images are usually stored (indeed, perhaps this is how our minds memorize three-dimensional images), in which case the angle between each view need be no finer than the minimum difference in projection angle needed to render two views of such an array distinct.

As physically perfect three-dimensional images, holograms have no flaws between views, and proponents sometimes unwisely claim that all else is mere compromise. But as the previous section demonstrated, a hologram is no different to the eye from an autostereoscopic display where the direction of view is controlled by diffraction, so holograms will also subtend a measurable angle between views, which, although too fine to see, will be finite. There will therefore also be a calculable depth of field, even for a hologram.

It is by relating the depth of field and angle per view between each pixellation scheme that the resolutions of differently pixellated images can be matched, and since one so often needs to display images of one format on a display of another, this section aims to formulate these relationships.

Dealing first with a cubic volumetric array, geometric optics is sufficient to determine the angle through which a video camera must move before its image of the array is substantially changed. Starting the camera far from the array but with sufficient magnification that each pixel at the front of the array maps onto one in the video camera, there will be a certain sideways distance through which the camera must move before the column of pixels at one side of the rear of the array map onto a fresh column of video camera pixels.

Fig. 11 shows that the angle $\Delta\theta$ subtended by this distance to the center front of the array equals the width

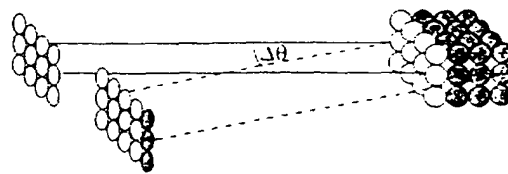


Fig. 11. Two views of a cubic array are formed by parallel projection. The minimum angle ($\Delta\theta$) between the views required for there to be a distinct difference in view content is that required to make one column of rear pixels fully visible.

of one array pixel Δx divided by the depth z of the array

$$\Delta\theta = \frac{\Delta x}{z}. \quad (1)$$

If the width of each pixel equals its depth and the array is n_z pixels deep, it follows that

$$\Delta\theta = \frac{1}{n_z}. \quad (2)$$

So the effective angle between views of a volumetric display is the reciprocal of the number of depth pixels, and the angle subtended by each view on an autostereoscopic display must equal this if it is to show a flawless image of equivalent depth. This means, for example, that the 3-D equivalent of a VGA image comprising 640 by 480 pixels will need approximately 480 views in azimuth if it is to represent an array as deep as it is high over a field of view of 60° (equal to about one radian).

Volumetric displays usually can image only finite depths, but in principle, autostereoscopic and holographic displays can act as windows into a three-dimensional environment. If the environment is effectively infinitely deep, comprising, for example, an object with mountains in the background, must the angle between views be infinitesimal?

The mistaken assumption in this question is that views of such an environment will be formed by parallel projection, i.e., to assume that views are formed by imaginary cameras far from the scene. In reality, the projections the cameras form will be not parallel but perspective, and each camera will be able to resolve fine resolution at a close distance but coarse resolution far away. The smallest object that can be resolved at any distance from the camera is equal to the width of view visible at that distance divided by the number of pixels per line in the camera. Rather than a uniform cubic array, a more appropriate test image is an array of volumetric pixels (voxels) in which the voxel dimension is proportional to the distance of the voxel from the camera, i.e., a distorted cubic array (Fig. 12).

Through what angle can the direction of projection be rotated before the projected image changes? If one rotates about the frontal center of the cubic array, the limit on rotation without change is set once again by the rear voxels of the array. There will have been an unambiguous change in the projected image once the direction of projection has been changed sufficient to translate the image of the rear voxels by one voxel diameter. Simple geometry shows that as the depth of the cubic array tends to infinity, this angle

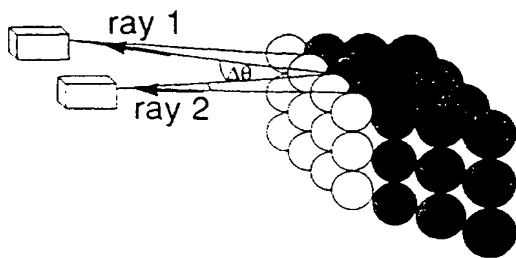


Fig. 12. In the left-hand video camera, only the front pixels of the distorted cubic array are visible, while in the right-hand camera, one column of the rear pixels has become visible. As the depth of the distorted cubic array tends to infinity, RAY 1 and RAY 2 tend toward parallel, so the angle between adjacent views tends to the angle between adjacent pixels.

equals the angle subtended by two voxels at the rear of the array to the camera.

If each pixel in the image plane of the central camera is mapped to a voxel at the rear of the distorted cubic array, then the angle through which the camera can be rotated before the image changes equals the angle between its aperture and two adjacent pixels in its image plane. It follows that in order to televise a pixellated three-dimensional image of a scene whose depth tends toward infinity, the angle $\Delta\theta$ between adjacent cameras should equal the cameras' field of view α divided by the number of pixels per line n_x

$$\Delta\theta = \frac{\alpha}{n_x}. \quad (3)$$

If such an image is to be accurately reproduced on an autostereoscopic display, then the angle at which rays from the edges of the display's screen converge should equal the field of view of the cameras. This is so that if the display is substituted for the original scene, the image recorded by the cameras is unchanged. Equation (3) therefore sets the angle between views on the display, so if a display with a field of view of 60° has VGA resolution views and is to act as an infinitely deep 3-D window, it needs approximately 640 views in azimuth.

While these translations between volumetric and autostereoscopic pixellation are correct geometrically, the angle per view of an autostereoscopic display is limited, and that of a holographic display is determined by the laws of diffraction. The angle between views $\Delta\theta$ on an autostereoscopic display cannot be less than the angular divergence $\delta\theta$ of the rays that constitute each view, which is approximately determined by the wavelength λ and the pixel diameter Δx according to the law of diffraction [65]

$$\delta\theta = \frac{\lambda}{\Delta x}. \quad (4)$$

The same approximate result can be derived in a more indirect manner from the gain-aperture relationship [66] that $\text{gain} = 4\pi/(\delta\theta)^2 = 4\pi(\text{Area})/\lambda^2$. It follows that

$$\Delta\theta \geq \frac{\lambda}{\Delta x}. \quad (5)$$

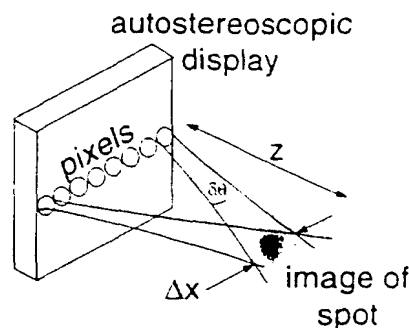


Fig. 13. An off-screen spot can be imaged by setting up rays to converge through it.

Continuing with the example of a display acting as an infinitely deep 3-D window with a field of view of 60° and VGA resolution views, the angle between views according to (3) was $1/640$ radians. So with red light ($\lambda = 630$ nm), (5) stipulates that the pixel size cannot be less than 0.4 mm. This is the approximate size of a pixel on a typical VGA monitor, so the restriction placed by diffraction on flawless autostereoscopically pixellated images is remarkably tight.

The depth of field (z) of an autostereoscopic display is the maximum distance above the screen at which light can be made to converge (as shown in Fig. 13) to form the image of a pixel of diameter Δx . By trigonometry, this distance approximately equals the pixel diameter divided by the angle of ray divergence

$$z \leq \frac{\Delta x}{\delta\theta}. \quad (6)$$

Equation (6) is essentially the same rule of geometry as (1) but referred to the coordinates of the display rather than those of the camera. Combining this with the law of diffraction given by (4) gives

$$z \leq \frac{(\Delta x)^2}{\lambda}. \quad (7)$$

In Section III, it was noted that although a three-dimensional image can be seen through a pair of raster scanning holes, the image will be slightly blurred. This blurring is caused by diffraction, and if the diameter of both holes is Δx , then (7) sets the maximum distance z between the scanning holes. Should the distance nevertheless be made greater than this, then diffraction through the second hole would filter detail rastered by the first scanning hole such that its effective size would increase to that allowed by (7).

Section III also noted that one can represent a single dimension of autostereoscopic pixellation on a diagram with coordinates of lateral angle Θ versus lateral position X . But the spatial angular frequency k_x of light waves of wavelength λ viewed in a plane intersecting the wave front at an angle Θ is given by

$$k_x = \frac{2\pi}{\lambda} \sin \Theta. \quad (8)$$

If autostereoscopic pixellation is represented instead by a plot of k_x versus X , then the dimensions of pixellation

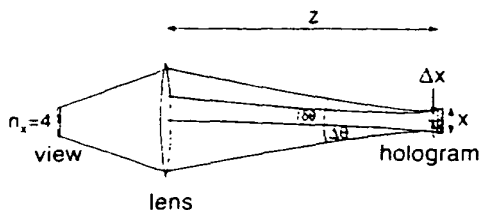


Fig. 14. The minimum angle per view ($\Delta\theta$) of a hologram equals the angle of divergence ($\Delta\theta$) of a single ray times the number of pixels (n_x) in a single row of the view.

are limited by combining (8) and (5) to get the classic expression

$$\Delta k_x \cdot \Delta x \geq 2\pi. \quad (9)$$

Depicting autostereoscopic pixellation as an array of independent subholograms of diameter Δx , (9) reaffirms that the minimum increment in spatial frequencies (Δk_x) that can be resolved by each subhologram is equal to 2π divided by its width (Δx).

Consider the effect of (7) on the three-dimensional equivalent of a high-resolution monitor, where pixellation can be as small as $90 \mu\text{m}$ [67]. The maximum depth of a cubic array would then be only 16 mm, so autostereoscopic systems are fundamentally inadequate for high-resolution 3-D images. Only holographic pixellation will suffice.

The angle between views on a hologram is also governed by the law of diffraction but in a slightly different way. The minimum divergence of any ray is determined not by the diameter of a pixel but, as shown in Fig. 14, by the diameter x of the whole hologram because the width of the wave fronts comprising the ray is ultimately limited by the edges of the hologram

$$\delta\theta = \frac{\lambda}{x}. \quad (10)$$

If a lens in the far field is only big enough to capture this ray, the image it forms will comprise a single spot of light: this does not constitute a view. A lens large enough to capture two rays will form what is in effect an image comprised of two pixels. Therefore, to form an image comprised of n_x pixels per line, the minimum angle of view that the lens must subtend is n_x times the minimum ray divergence

$$\Delta\theta \geq \frac{n_x \lambda}{x}. \quad (11)$$

The lens will have to be moved through this entire angle before it forms a new and independent view, and if it is merely moved part of the way between, then it will form a superposition of the adjacent views.

The choice of n_x in this instance is somewhat arbitrary, but once made, then the smallest pixel diameter (Δx) that can be resolved on the hologram is by simple geometry

$$\Delta x = \frac{x}{n_x}. \quad (12)$$

Combining this with (11) gives

$$\Delta\theta \geq \frac{\lambda}{\Delta x}. \quad (13)$$

This is exactly the same as (5), so if a hologram's pixel size is defined to be the same as that of a diffraction-limited autostereoscopic image, both have the same angle per view, and therefore both have the same information content for the same quality of three-dimensional image. Since it has already been shown that a flawless autostereoscopic image with equivalent resolution and size to a conventional VGA monitor is at the diffraction limit, it follows that under typical conditions, a flawless autostereoscopic image contains no less information than a hologram.

Diffraction effects will not matter with large displays, nor when cameras are imaging large scenes. But if high-resolution images are being formed of small-scale phenomena—as would be required, for example, in 3-D keyhole surgery—diffraction effects in the 3-D camera will need to be considered and will obey rules similar to those given above.

The depth of field of a hologram is found by combining (6) and (10)

$$z \leq \frac{x \cdot \Delta x}{\lambda}. \quad (14)$$

So a hologram of width $x = 20 \text{ cm}$, for example, illuminated by the light of wavelength $\lambda = 500 \text{ nm}$, could in theory project a spot of diameter $\Delta x = 100 \mu\text{m}$ up to 40 m from its surface. Closer to its surface, the smallest spot that a hologram can project is equal to its resolution, which in principle can be as small as one wavelength of light. Table 1 summarizes these relationships.

One class of autostereoscopic displays is inherently exempted from the restrictions described in this section—those that use coherent light or bulk optics. These are a special case because light is coherent across the optical wave front. So although if all the screen is opaque except for one pixel then light will diffract as with incoherent autostereoscopy, if several adjacent pixels on a bulk optical display are transparent, then light will diffract less. Indeed, one can imagine writing a zone plate on such a display in order to cause an off-screen pixel to come into focus somewhere above the screen. Displays of this kind make possible an intermediary between autostereoscopic and holographic pixellation by combining them.

V. HYBRID PIXELLATION

The ideal hologram has a minimum pixel size that is related to the maximum angle of view by (13). With the $10\text{-}\mu\text{m}$ pixel width typical of present spatial light modulators operating on $\sim 500\text{-nm}$ wavelengths at the center of the visual spectrum, angles of view are limited to approximately $1/20$ radian. So it is proposed to combine autostereoscopic and holographic pixellation into a hybrid scheme that interchanges the concept of projection of view over a range of angles, as already discussed for

Table 1 The Angle Per View ($\Delta\theta$) and Depth (z) of the Three Pixellation Schemes Can Be Related By the Width of the Image (r), the Depth of the Image in Pixels (n_z), the Width of the Image in Pixels (n_x), the Field of View of the Image (α), the Pixel Size (Δr), and the Wavelength (λ)

	Distorted volumetric	Cartesian volumetric	Autostereo	Holographic
angle per view	$\Delta\theta = \frac{\alpha}{n_z}$	$\Delta\theta = \frac{1}{n_z}$	$\Delta\theta \geq \frac{\lambda}{\Delta x}$	$\Delta\theta \geq \frac{n_z \lambda}{x}$
Depth	∞	$z = n_z \Delta z$	$z \leq \frac{(\Delta x)^2}{\lambda}$	$z \leq \frac{x \cdot \Delta x}{\lambda}$

autostereoscopic systems, to the projection of a narrow-angle hologram over a range of views.

The approximate algebra that explains the details of this concept is as follows. The light from a hologram is centered around the central wavelength and, for the purposes of this discussion, can be taken to be at one wavelength with a wave propagation vector whose magnitude k equals $2\pi/\lambda$. A one-dimensional range of angles will be considered, with k_z giving the axial component of the wave propagation vector and k_x giving the lateral component. For the purposes of this discussion, the approximation is made that $k_z \approx 2\pi/\lambda$ and the lateral angle of projection $\Theta \approx \lambda k_x / 2\pi$. Ideally, then, one wishes to have the hologram projection into a range of angles, say $-\frac{1}{2}\theta$ to $+\frac{1}{2}\theta$, with these angles then corresponding to values of $k_x = -\kappa$ to $+\kappa$, respectively, where $\lambda\kappa/2\pi = \frac{1}{2}\theta$. If the complex field amplitude is $\tilde{E}(k_x)$ corresponding to a direction determined by k_x , then the near-field amplitude in the plane of the spatial light modulator is, say, $E(X)$, where from the Fourier transform we get

$$E(X) = \int_{-\kappa}^{+\kappa} \tilde{E}(k_x) \exp(jk_x X) \frac{dk_x}{2\pi}. \quad (15)$$

Now we recognize that this could be accomplished by a superposition of N narrow-angle holograms, each giving a total angle $\Delta\theta = \lambda 2\kappa / N 2\pi$. Then, writing $\Delta\kappa = 2\kappa/N$, one may split the hologram

$$E(X) = \sum_{m=1}^N E_m(X) \quad (16)$$

where $E_m(X)$ expands thus:

$$E(X) = \sum_{m=1}^N \int_{-\kappa+(m-1)\Delta\kappa}^{-\kappa+m\Delta\kappa} \tilde{E}(k_x) \exp(jk_x X) \frac{dk_x}{2\pi}. \quad (17)$$

The variable k_x is changed for each range

$$k_x = k'_x + [-\kappa + (m - \frac{1}{2})\Delta\kappa] \quad (18)$$

to give a narrower range for k'_x than for k_x . This lets us write

$$E(X) = \sum_{m=1}^N \exp\{j[-\kappa + (m - \frac{1}{2})\Delta\kappa]X\} E_m(X) \quad (19)$$

where

$$E_m(X) = \int_{-(1/2)\Delta\kappa}^{+(1/2)\Delta\kappa} \tilde{E}_m(k'_x) \exp(jk'_x X) \frac{dk'_x}{2\pi}. \quad (20)$$

The exponential expression in (19) is the Fourier transform of $\delta(-\kappa + (m - \frac{1}{2})\Delta\kappa)$, so letting $FT\{\}$ denote the operation of taking a Fourier transform, we can write

$$E(X) = \sum_{m=1}^N E_m(X) FT\{\delta[k + \kappa - (m - \frac{1}{2})\Delta\kappa]\}. \quad (21)$$

The values of $E_m(X)$ give the required near-field pattern on the spatial light modulator where the pixel width can now be as big as $N\lambda/\theta$ but the m th holographic view is projected at an angle approximately equal to $\lambda[-\kappa + (m - \frac{1}{2})\Delta\kappa]/2\pi$. Since the operation of a lens on light traveling from one focal plane to the other can be represented by a Fourier transform [68], (21) indicates that the projection of each holographic view can be achieved by putting a lens behind the spatial light modulator and placing a point source of light representing the impulse in the focal plane of the lens. Of course, there is no such thing as a point source of light, and it is assumed that the summation of far-field intensities will be carried out time sequentially such as to give a summation of far-field intensities rather than the summation of complex amplitudes specified by the algebra. Furthermore, the details of the algebra will need alteration for realistically large angles, although the concept remains and will not materially change. Nevertheless, given the insensitivity of the eye to phase, (21) indicates that by exchanging the views displayed on the system described

in Section I with a series of holograms, and by replacing the scanning spot source of light with something as close as possible to a point source moving in discrete steps, the result will be a fault-free three-dimensional image.

There is an important issue for holograms produced by simply modulating the intensity of the light by a spatial light modulator. Such a hologram is often referred to as a binary phase hologram and produces symmetric patterns for $\pm k'_x$. One can see that simply modulating an intensity pattern with a spatial light modulator will not be effective, as it will only be able to produce a series of symmetric patterns about each projected angle, even though the eye is insensitive to phase. It is consequently envisaged that a spatial light modulator designed to modulate four different phase levels will be required to remove this symmetry [69].

The field of view θ of a hybrid three-dimensional image is the number of views n_θ times the angle per view governed by (13)

$$\theta = n_\theta \frac{\lambda}{\Delta x}. \quad (22)$$

Assuming a flicker rate of 50 Hz, the frame rate of the spatial light modulator must be at least $50 n_\theta$ Hz, and the limit of spatial periodicity will be the reciprocal of the pixel spacing (Δx). Defining space-time periodicity to equal the product of frame rate and the reciprocal of pixel spacing, (22) shows that for the display of a hybrid three-dimensional image with field of view θ , the space-time periodicity must be greater than fifty times the field of view divided by the wavelength. For a one-radian field of view in azimuth with a wavelength of 500 nm, the space-time periodicity should approximately equal $1 \text{ Mbs}^{-1} \text{cm}^{-1}$, well above the $5\text{-kbs}^{-1} \text{cm}^{-1}$ capabilities of large high-resolution liquid crystal displays [67].

It might seem premature to be considering the ultimate resolution of three-dimensional video images when present resolutions are so much lower. But while the resolution of spatially multiplexed and time-multiplexed autostereoscopic displays is limited by spatial resolution and frame rate, respectively, what the hybrid approach offers is the ability to interchange spatial and temporal periodicity. It is then the product of spatial and temporal periodicity that determines what three-dimensional resolution a device makes possible, and devices already exist with the space-time periodicities necessary for high-resolution three-dimensional images.

VI. ADVANCED 3-D DISPLAYS

If liquid crystal displays lack the space-time periodicity needed for hybrid pixellation, one device stands out for its lack of complexity and high space-time periodicity: the light valve (also known as an optically addressable spatial light modulator). Section II notes that video holograms have already been screened by optically addressing such a device with a cathode ray tube, but the field of view was narrow. By combining the frame rate of the latest light valves [70] with techniques for phase modulation [71], [72],

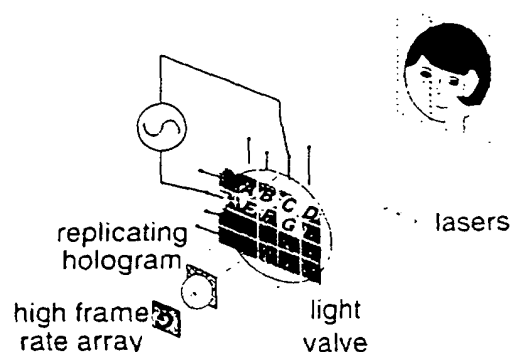


Fig. 15. An autostereoscopic/holographic display with a wide field of view can be made by time sequentially illuminating a high-resolution liquid crystal display, which can be assembled from a light valve and a high-frame-rate array.

it should be possible to effect hybrid pixellation so as to obtain considerably wider fields of view. The rate at which data must be fed to such a device in order for it to operate effectively equals the number of pixels in the device times its frame rate, called the space-bandwidth product. High-frame-rate arrays now have the space-bandwidth product needed to address light valves over large areas but tend to have fewer pixels and higher frame rates than light valves. Fig. 15 shows how the light valve can be addressed despite this by multiplexing the image of an array across its rear. The great advantage of this approach is that it removes from the screen (i.e., the light valve) the two most expensive items: the active matrix transistors and the connector array. One is then left with a screen that may be large but is uncomplicated and a small video projector that may be complicated but is not large. Both devices are therefore potentially cheap, and it is encouraging to note that arguably, it is just this division of size and complexity between phosphor screen and electron gun that made it so economic to manufacture cathode ray tubes.

While hybrid pixellation provides for flawless three-dimensional images, it remains unclear that users object to minor flaws, and autostereoscopic pixellation would certainly be the simplest to implement on such a device if it were fast enough. But the frame rate of the latest light valves seems to be limited to approximately 2 kHz by the resistor/capacitor time constant of the amorphous silicon. Dividing this by three for color and by 60 for flicker, one might get 30 views, but if these views have a typical 640 pixels per line and are taken by cameras with a view of half a radian (approximately 30°), then according to (2), for a flawless image the angle per view should be $1/1280$ radians and the field of view of the device would be less than $1/40$ radians (approximately 1.5°). Of course, 30 views at the 1° per view that seems acceptable for the first generation of video three-dimensional images would result in a satisfactory field of view. But the optically addressed system is not a flat panel, and with autostereoscopic pixellation it would produce a three-dimensional image little better than the flat-panel active matrix liquid crystal display. The extra cost of the latter will eventually depend on how many get

made, but in large quantities it might be low enough to win over optical addressing.

A typical light valve can resolve down to $10\text{ }\mu\text{m}$, which with a frame rate of 2 kHz gives a space-time periodicity of $2 \times 10^{13}\text{ m}^{-2}\text{s}^{-1}$. After dividing by three for color and 60 for flicker, one can estimate the solid angle available for viewing by multiplying by λ^2 , equal approximately to $(0.5 \times 10^{-6})^2\text{ m}^2$. The result is a solid angle of view of 0.025, equivalent to a viewing zone of, say, 30° in azimuth by 3° in elevation.

Light valves are likely to be able to do better than this. Frame rates of 5 kHz have been reported [73] at the penalty of intense illumination (and a bistable liquid crystal), as have spatial resolutions of $5\text{ }\mu\text{m}$. But before drawing optimistic conclusions, one should consider the problem of writing data to these devices at rates approaching 400 GHz for a 16 by 12 cm screen.

An optical fiber is capable of transmitting data at such rates, and a simple method of scanning its output would be a tremendous prize both for displays and telecommunications. But existing acoustooptic devices can barely scan at 1 MHz , and optical amplifier arrays remain rather elementary. It was research into photonics that led to fast-switching light valves, and it is research into photonics that is producing some of the more promising ways of addressing them. If the addressing problem is simplified by requiring an image that is three dimensional only in azimuth, then for a 16-cm -wide screen with 240 interlaced lines, the data rate reduces to (frame rate \times lateral resolution \times width \times number of lines) $= (2000 \times 10^5 \times 0.16 \times 240/2) \approx 4\text{ GHz}$. This brings the data rate within the range of existing devices, and five stand out: acoustooptic holograms, cathode ray tubes, laser diode arrays, ferroelectric arrays, and micromirror arrays.

Acoustooptic holograms have a successful history but are limited by the speed of sound in acoustooptic materials, which at 5 km/s restricts data rates to approximately 10 GHz for an optical wavelength of half a micrometer. In practice, even these rates are difficult because of the attenuation at high frequencies mentioned in Section II.

Cathode ray tubes can be electrostatically scanned at megahertz line rates, and, providing the deflection angle is narrow and the beam intensity is not too high, the spot size can be kept to a diameter of a few micrometers. But it is difficult to make such a small spot bright without defocusing, and a way must be found of fully modulating the intensity of a dense electron beam at more than 1 GHz . While these challenges are not insuperable, they remain challenges.

Laser diode and other arrays work by demultiplexing the input to a sufficient resolution that raster scanning is either not required or need be no faster than can be executed by a liquid crystal hologram. An 18×1 laser diode array has been operated at $18 \times 1\text{ GHz}$, and 256×256 arrays have been fabricated, offering the tantalizing prospect of space-bandwidth products far in excess of any alternative.

Ferroelectric arrays are fast-switching liquid crystal displays where the active matrix transistors are etched in a

silicon integrated circuit. A 320×240 array with a potential frame rate of 20 kHz has been demonstrated [74], offering a space-bandwidth product of 1.5 GHz , which is getting close to that needed for a 16-cm -wide screen.

Micromirror arrays [75] have the advantage of being comprised entirely of silicon, although they require a more intricate lithography. Nevertheless, arrays of 2048×1152 pixels potentially offer a space-bandwidth product of 5.8 GHz . Details of circuitry aside, this offers the potential for a screen more than eight inches wide, and if three such devices were operated in parallel (which is how they are configured for high-definition 2-D projection), then one could hope for better quality still. But once again, optimistic conclusions are inappropriate, in this case because these devices merely convert data from an electronic form to an optical one; a source of data is still required.

Whatever the capabilities of optical fiber, it seems highly probable that three-dimensional images will be compressed. The convention at present is that displays are connected to a video driver by a cable and that any decompression is effected by the video driver. But the data rates of raw three-dimensional video are so high that it seems pointless to decompress the signal remote from the display merely then to be presented with the challenge of transmitting a raw signal. Rather, the decompression should take place as close to the addressing device as possible (perhaps even within the addressing device), and it is convenient that both micromirror and ferroelectric arrays are mounted on carriers that plug directly into a printed circuit board. The complexity and output data rate of existing interfaces for three-dimensional video suggests that the decompression machine will have computational power comparable to that of a typical computer, and with the current trend for the display to dominate the cost of a computing system, it must be questioned whether there continues to be any advantage in going to the effort of separating the computer from the display.

This section has brought the paper to a conclusion by attempting to demonstrate in some detail that it is practicable with existing technology to display a medium-sized color three-dimensional video image with no moving parts, an adequate field of view, and no flaw lines. Three-dimensional video is not some remote or esoteric prospect: it is a viable, analytic technology, and its development, like that for two-dimensional video, will depend on further progress in the three fundamentals of display technology—spatial demultiplexing, screen space-bandwidth product, and low cost per unit screen area.

VII. CONCLUSIONS

Video three-dimensional images can be pixellated in three ways: volumetric, holographic, and autostereoscopic. While volumetric images use bandwidth efficiently to give all-round viewing and holographic displays have high resolution, autostereoscopic displays image opaque objects with the wide fields of view needed for most applications.

Autostereoscopic displays that track viewers' heads offer the prospect of greatly reduced data rates, but multiple-view

autostereoscopy avoids the need for machine intelligence. The latest such displays time-multiplex views to get the 1° view spacing that seems adequate for the first generation of three-dimensional displays.

Although acceptable in the short term, the images on autostereoscopic displays with 1° per view are flawed and may come to irritate. For true three-dimensional images, the angle per view must be approximately 1/10° for an image 640 pixels wide. At this spacing, even an autostereoscopic display with pixel diameter as big as 0.5 mm will be diffraction limited, and its data content no less than that of a hologram. Holograms have greater depths of field than autostereoscopic images and much greater resolution, and are virtually the only option for pixel sizes finer than 0.5 mm.

Holographic and time-multiplexed autostereoscopic pixellation schemes can be combined to give a hybrid that has the virtues of both. A sequentially illuminated holographic display has the same data content, resolution, and depth as a hologram but the field of view of an autostereoscopic display. In principle, all that is needed is a liquid crystal display with a space-time periodicity of the order of $1 \text{ Mb-s}^{-1} \text{ cm}^{-1}$, but this is impractical over large areas at low cost.

Faced with the demand for high space-bandwidth products, the optical communications industry has developed light valves and high-frame-rate arrays sufficient to get the requisite space-time periodicities over large areas. Light valves are simple enough to operate over screen-sized areas at low device cost, and the arrays provide a way of spatially distributing data across the light valve that need not be expensive, provided they are small. Projecting a small array onto a large light valve therefore gives a display that is cheap and has high resolution for the same reasons that the cathode ray tube does.

The high-frame-rate array should be as close as possible to the electronics that decompress the three-dimensional image in order to minimize high-data-rate connections. The computational power of the decompression electronics will be comparable to that of most computers, and it is not unlikely that the computer and display systems will therefore come to merge.

The progress of two-dimensional video has since its invention been one of steady evolution toward increasing resolution and size, drawing on parallel advances in telecommunications. While the display of video three-dimensional images may seem revolutionary, this paper has sought to show that the pixellation and display optics are not unduly sophisticated and that the remaining challenges are the same as for two-dimensional video: an increase in the screen's space-bandwidth product, an increase in the rate at which data can be physically distributed across the screen, and the attainment of both in a single system without great complexity of manufacture. Photonic components developed for optical telecommunications already meet the requirements for three-dimensional video, and the two technologies are likely to continue to interact to their mutual benefit.

ACKNOWLEDGMENT

The author wishes to thank the following for useful discussions: L. W. K. Yim on light valves, S. R. Lang on head tracking, W. A. Crossland on high-bandwidth photonics, and J. E. Carroll for his comments on this article.

REFERENCES

- [1] J. E. Wickham, "Minimally invasive surgery future developments," *Br. Medical J.*, vol. 308, pp. 193-196, Jan. 1994.
- [2] T. Motoki, H. Isono, and I. Yuyama, "Recent status of 3-dimensional television research," *Proc. IEEE*, vol. 83, pp. 1009-1021, July 1995.
- [3] A. C. Traub, "Three-dimensional display," U.S. Patent 3,493,290, Jan. 14, 1966.
- [4] P. H. Mills, H. Fuchs, and S. M. Pizer, "High speed interaction on a vibrating mirror 3D display," in *Proc. SPIE*, vol. 507, 1994, pp. 93-101.
- [5] C. C. Tsao, and J. S. Chen, "Moving screen projection: A new approach for volumetric three-dimensional display," in *Proc. SPIE*, vol. 2650, 1996, pp. 254-264.
- [6] K. Kameyama, K. Ohtomi, and Y. Fukui, "Interactive volume scanning 3D display with an optical relay system and multidimensional input devices," in *Proc. SPIE*, vol. 2915, 1993, pp. 12-20.
- [7] H. Yamada, K. Yamamoto, M. Matsushita, J. Koyama, and K. Miyaji, "3D display using laser and moving screen" in *Proc. Japan Display '89, Society for Information Display*, Kyoto, Japan, 1989, pp. 630-633.
- [8] K. Kameyama, K. Ohtomi, and Y. Fukui, "Interactive volume scanning 3-D display with an optical relay system and multidimensional input devices," in *Proc. SPIE*, vol. 1915, 1993, pp. 12-20.
- [9] D. G. Jansson, and R. P. Kosowsky, "Display of moving volumetric images," in *Proc. SPIE*, vol. 507, 1984, pp. 82-92.
- [10] M. E. Lasher, P. Soltan, W. J. Dahlke, N. Acantilado, and M. McDonald, "Laser-projected 3D volumetric displays," in *Proc. SPIE*, vol. 2650, 1996, pp. 285-295.
- [11] D. Bahr, K. Langhans, M. Gerken, C. Vogt, D. Bezecky and D. Homann, "FELIX: A volumetric 3D laser display," in *Proc. SPIE*, vol. 2650, 1996, pp. 265-273.
- [12] R. D. Williams and F. Garcia, "A real-time autostereoscopic multiplanar 3D display system," in *Proc. SID Int. Symp.*, Anaheim, CA, May 24-26, 1988, pp. 91-94.
- [13] B. G. Blundell, A. J. Schwarz, and D. K. Horrell, "Cathode-ray sphere—A prototype system to display volumetric 3-dimensional images" *Opt. Eng.*, vol. 33, no. 1, pp. 180-186, 1994.
- [14] E. A. Downing, L. Hesselink, R. M. Macfarlane, and C. P. Barty, "Solid-state three-dimensional computer display," *Proc. Conf.—Lasers and Electro-Optics Society Annual Meeting*, Anaheim, CA, vol. 8, 1994, pp. 6-7.
- [15] I. I. Kim, E. J. Korevaar, and H. Hakakha, "Three-dimensional volumetric display in rubidium vapor," in *Proc. SPIE*, vol. 2650, 1996, pp. 2274-2284.
- [16] T. Hattori, D. F. McAllister, and S. Sakuma, "Spatial modulation display using spatial light modulators," *Opt. Eng.*, vol. 31, pp. 350-352, Feb. 1992.
- [17] K. Higuchi, K. Ishii, J. Ishikawa, and S. Hiyama, "Experimental holographic movie IV: The projection-type display system using a retro-directive screen," in *Proc. SPIE*, vol. 2406, 1995, pp. 20-26.
- [18] H. Katsuma and K. Sato, "Electronic display system using LCD, laser-diode, and holography camera," in *Proc. SPIE*, vol. 1914, 1993, pp. 212-218.
- [19] K. Maeno, N. Fukaya, O. Nishikawa, K. Sato, and T. Honda, "Electro-holographic display using 15mega pixels LCD," in *Proc. SPIE*, vol. 2652, 1996, pp. 15-23.
- [20] N. Hashimoto and S. Morokawa, "Motion-picture holography using liquid-crystal television spatial light modulators," in *Proc. SID Int. Symp.*, Orlando, FL, vol. 26, 1995, pp. 847-850.
- [21] M. W. Thie, J. Lukins, and D. A. Gregory, "Optically addressed SLM-based holographic display," in *Proc. SPIE*, vol. 2488, 1995, pp. 408-416.

- [22] H. Farhoosh, Y. Fainman, K. Urquhart, and S. H. Lee, "Real-time display of 3-D computer-data using computer generated holograms," in *Proc. SPIE*, vol. 1052, 1989, pp. 172-176.
- [23] M. Lucente, R. Pappu, C. J. Sparrell, and S. A. Benton, "Progress in holographic video with the acoustooptical modulator display," in *Proc. SPIE*, vol. 2577, 1995, pp. 2-7.
- [24] J. S. Kollin, "Time multiplexed auto stereoscopic three dimensional imaging system," U.S. Patent 4853769, June 16, 1987.
- [25] J. Y. Son, S. A. Shestak, S. K. Lee, and H. W. Jeon, "Pulsed laser holographic video," in *Proc. SPIE*, vol. 2652, 1996, pp. 24-28.
- [26] T. Yamazaki, K. Kamijo, and S. Fukuzumi, "Quantitative evaluation of visual fatigue encountered in viewing stereoscopic 3D display—Near point distance and visual evoked potential study," in *Proc. Japan Display '89, Society for Information Display*, Kyoto, Japan, 1989, pp. 606-609.
- [27] D. J. Trayner and E. Orr, "Autostereoscopic display using holographic optical elements," in *Proc. SPIE*, vol. 2653, 1996, pp. 65-74.
- [28] P. V. Harman, "Autostereoscopic display system," in *Proc. SPIE*, vol. 2653, 1995, pp. 56-64.
- [29] D. Ezra, "Look, no glasses," *Inst. Electr. Eng. Rev.*, vol. 42, pp. 187-189, Sept. 1996.
- [30] N. Tetsutani, K. Omura, and F. Kishino, "Wide-screen autostereoscopic display system employing head-position tracking," *Opt. Eng.*, vol. 33, pp. 3690-3697, Nov. 1994.
- [31] H. Imai, M. Imai, Y. Ogura, and K. Kubota, "Eye-position tracking stereoscopic display using image shifting optics," in *Proc. SPIE*, vol. 2653, 1996, pp. 49-55.
- [32] L. McMillan, G. Bishop, "Head-tracked stereoscopic display using image warping," in *Proc. SPIE*, vol. 2409, 1995, pp. 21-30.
- [33] A. Katayama, K. Tanaka, T. Oshino, and H. Tamura, "Viewpoint-dependent stereoscopic display using interpolation of multiviewpoint images," in *Proc. SPIE*, vol. 2409, 1995, pp. 11-20.
- [34] T. Hattori, "On the wall stereoscopic liquid crystal display," in *Proc. SPIE*, vol. 2409, 1995, pp. 41-47.
- [35] S. Omori, J. Suzuki, S. Sakuma, "Stereoscopic display system using backlight distribution," in *Proc. SID Int. Symp.*, Orlando, FL, vol. 26, 1995, pp. 855-858.
- [36] K. Mase, Y. Watanabe, Y. Suenaga, "A realtime head motion detection system," in *Proc. SPIE*, vol. 1260, 1990, pp. 262-269.
- [37] K. C. Yow and R. Cipolla, "A probabilistic framework for perceptual grouping of features for human face detection," in *Proc. 2nd Int. Conf. Automatic Face and Gesture Recognition*, Killington, VT, 1996, pp. 16-21.
- [38] H. Isono, M. Yasuda, H. Kusaka, and T. Morita, "3D flat-panel displays without glasses," in *Proc. Japan Display '89, Society for Information Display*, Kyoto, Japan, 1989, pp. 626-629.
- [39] M. Brewin, M. Forman, and N. A. Davies, "Electronic capture and display of full-parallax 3D images," in *Proc. SPIE*, vol. 2409, 1995, pp. 118-124.
- [40] J. Hamasaki, M. Okada, S. Utsunomiya, S. Uematsu, and O. Takeuchi, "Autostereoscopic 3D TV on a CRT," presented at the SID International Symposium, Anaheim, CA, May 9, 1991.
- [41] J. Guichard and A. Poirier, "An experiment in three-dimensional television," *Radiodiffusion-Telev* (France), vol. 20, 1986, pp. 23-29.
- [42] C. van Berkel, A. R. Franklin, and J. R. Mansell, "Design and applications of multiview 3D-LCD," in *Proc. Euro Display '96, Society for Information Display*, Birmingham, U.K., 1996, pp. 109-112.
- [43] K. Sakamoto, M. Okamoto, H. Ueda, H. Takahashi, and E. Shimizu, "Real-time 3-D color display using a holographic optical element," in *Proc. SPIE*, vol. 2652, 1996, pp. 124-131.
- [44] G. P. Nordin, J. H. Kulick, M. Jones, P. Nasiatka, R. G. Lindquist, and S. T. Kowel, "Demonstration of a novel three-dimensional autostereoscopic display," *Opt. Lett.*, vol. 19, pp. 901-903, June 1994.
- [45] T. Toda, S. Takahashi, and F. Iwata, "Three-dimensional (3D) video system using grating image," in *Proc. SPIE*, vol. 2652, 1996, pp. 54-61.
- [46] R. Borner, "Autostereoscopic 3D-imaging by front and rear projection and on flat-panel displays," *Displays*, vol. 14, pp. 39-46, 1993.
- [47] G. Bader, E. Lueder, and J. Fuhrmann, "An autostereoscopic real-time 3D display system," in *Proc. Euro Display '96, Society for Information Display*, 1996, pp. 101-104.
- [48] Y. Kajiki, H. Yoshikawa, and T. Honda, "3D display with focused light array," in *Proc. SPIE*, vol. 2652, 1996, pp. 106-116.
- [49] G. B. Meacham, "Autostereoscopic displays—past and future," in *Proc. SPIE*, vol. 624, 1986, pp. 90-101.
- [50] R. J. Felix, "Multiplex video display," in *Proc. SPIE*, vol. 2176, 1994, pp. 50-56.
- [51] J. B. Eichenlaub, "Autostereoscopic display with illuminating lines and light valve," U.S. Patent 4717949, Jan. 5, 1988.
- [52] J. B. Eichenlaub, D. Hollands, and J. M. Hutchins, "A prototype flat plane hologram-like display that produces multiple perspective views at full resolution," in *Proc. SPIE*, vol. 2409, 1995, pp. 102-112.
- [53] A. R. Travis, "Autostereoscopic 3-D display," *Appl. Opt.*, vol. 29, pp. 4341-4343, Oct. 1990.
- [54] H. Isono, M. Yasuda, and H. Sasazawa, "Multi-viewpoint 3D display with time-divided backlighting system," *Electron. Commun. Jpn. 2, Electron. (USA)*, vol. 76, no. 7, pp. 77-84, July 1993.
- [55] K. Nito, T. Fujioka, N. Kataoka, and A. Yasuda, "TFT-driven monostable ferroelectric liquid crystal display with wide viewing angle and fast response times," in *Proc. AM-LCD '94, Society for Information Display*, Tokyo, Japan, Nov. 1994, pp. 48-51.
- [56] L. Noble, "Use of lenses to enhance depth perception," in *Proc. SPIE*, vol. 761, 1987, pp. 126-128.
- [57] R. B. Collender, "3D television, movies and computer graphics without glasses," *IEEE Trans. Consumer Electron.*, vol. CE-32, pp. 56-61, 1986.
- [58] H. B. Tilton, "An autostereoscopic CRT display," in *Proc. SPIE*, vol. 120, 1977, pp. 68-72.
- [59] R. G. Batchko, "Three-hundred-sixty degree electroholographic stereogram and volumetric display system," in *Proc. SPIE*, vol. 2176, 1994, pp. 30-41.
- [60] A. R. Travis, S. R. Lang, J. R. Moore, and N. A. Dodgson, "Time-multiplexed three-dimensional video display," *J. SID*, vol. 3, pp. 203-205, 1995.
- [61] J. L. Baird, "Stereoscopic color television," *Wireless World*, vol. 48, pp. 31-32, Feb. 1942.
- [62] R. Herbert, "J. L. Baird's color television 1937-46," *Television: J. Royal Television Society*, pp. 24-29, Jan./Feb. 1990.
- [63] T. A. Theoharis, A. R. Travis, and N. E. Wiseman, "3D display: synthetic image generation and visual effect simulation," *Comput. Graph. Forum*, vol. 9, pp. 337-348, 1990.
- [64] S. Pastoor and K. Schenke, "Subjective assessments of the resolution of viewing directions in a multiviewpoint 3DTV system," in *Proc. Society for Information Display*, vol. 30, no. 3, 1989, pp. 217-223.
- [65] E. Hecht, *Optics*, 2nd ed. Reading, MA: Addison-Wesley, 1989, p. 419.
- [66] S. Ramo, J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communication Electronics*, 3rd ed. New York: Wiley, 1994, ch. 12, p. 666.
- [67] R. Martin, T. Chuang, H. Steemers, R. Allen, R. Fulks, S. Stubo, D. Lee, M. Young, J. Ho, M. Nguyen, W. Meuli, T. Fiske, R. Bruce, M. Thompson, M. Felton, and L. D. Silverstein, "A 6.3-Mpixel AMLCD," in *SID Int. Symp. Dig. Technical Papers*, Seattle, WA, May 18-20, 1993, vol. XXIV, pp. 704-707.
- [68] E. Hecht and A. Zajac, *Optics*. Reading, MA: Addison-Wesley, 1974, ch. 11, p. 412.
- [69] S. E. Broomfield, M. A. Neil, and E. G. Paige, "A 4-level, phase-only, spatial light modulator," *Electron. Lett.*, vol. 29, pp. 1661-1663, 1993.
- [70] L. W. Yim, A. B. Davey, and A. R. Travis, "Optically addressed spatial light modulators using the twisted smectic C' liquid crystal effect," *Ferroelectrics*, vol. 181, pp. 147-160, 1996.
- [71] T. D. Wilkinson and R. J. Mears, "Breaking symmetry in the binary phase only matched filter," *Opt. Commun.*, vol. 115, pp. 26-28, Mar. 1995.
- [72] M. A. Neil and E. G. Paige, "Breaking of inversion symmetry in 2-level, binary, Fourier holograms," in *Proc. Inst. Electr. Eng. Holographic Systems, Components and Applications*, Neuchatel, Switzerland, 1993, ch. 51, pp. 85-90.
- [73] F. Perennes, W. A. Crossland, D. Kozlowski, and Z. Y. Wu, "New reflective layer technologies for fast ferroelectric liquid

- crystal optically addressed spatial light modulators." *Ferroelectrics*, vol. 181, pp. 129-137, 1996.
- [74] T. D. Wilkinson, W. A. Crossland, T. Coker, A. B. Davey, M. Stanley, T. C. Yu, "The fast bitplane SLM: a new ferroelectric liquid crystal on silicon spatial light modulator," *Spatial Light Modulators*, Technical Dig. (Optical Society of America), pp. 149-150, 1997.
- [75] L. J. Hornbeck, "Digital light processing and MEMS: timely convergence for a bright future," presented at Micromachining and Microfabrication '95, Austin, TX, Oct. 1995.



A. R. L. Travis received the Ph.D. degree in optical fiber multiports for coherent detection from Cambridge University, U.K., in 1988.

Since 1988, he has been a Fellow of Clare College, Cambridge, and a Lecturer at Cambridge University. His research is on three-dimensional video.